

Army Intelligence Analysis and Interpretation: Assessing the Utility and Limitations of Computational Diagnostic Reasoning

Gerald M. Powell, Ph.D.

U.S. Army Research, Development and Engineering Command
Communications-Electronics Research, Development and Engineering Center
Intelligence and Information Warfare Directorate
AMSRD-CER-IW-BM
Fort Monmouth, NJ 07703

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE JUN 2004		2. REPORT TYPE		3. DATES COVERED 00-00-2004 to 00-00-2004	
4. TITLE AND SUBTITLE Army Intelligence Analysis and Interpretation: Assessing the Utility and Limitations of Computational Diagnostic Reasoning				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Army Research Development and Engineering Command, Communications-Electronics Research, Development and Engineering Center, AMSRD-CER-IW-BM, Fort Monmouth, NJ, 07703				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 11	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Army Intelligence Analysis and Interpretation: Exploring the Utility and Limitations of Computational Diagnostic Reasoning

Gerald M. Powell, Ph.D.

U.S. Army Research, Development and Engineering Command
Communications-Electronics Research, Development and Engineering Center
Intelligence and Information Warfare Directorate
AMSRD-CER-IW-BM
Fort Monmouth, NJ 07703

Abstract

The U.S. Army's Future Force is critically dependent on information superiority levels that will support timely, quality decision-making during high tempo operations. The Future Force is anticipated to produce unprecedented levels of data requiring analysis. Data fusion is considered by some as the potential solution to handling this severe data overload. The Joint Directors of Laboratories Data Fusion model categorizes data fusion-related functions at a high level of generality. For fusion in the Army, little has been published reflecting an elaboration of functionality on levels 2 and 3 of this model, both of which are viewed as critical elements of intelligence analysis and interpretation. Walsh (2002) offered a first-level decomposition of functionality for Level 2. Powell and Broome (2002) and Powell (2002) indicated the complex set of interrelationships among problems within and between fusion levels characterizing Army intelligence analysis and interpretation suggests a user-centric, holistic approach addressing fusion levels 1 through 5. The present paper characterizes selected key aspects of analysis and interpretation problems and processes based on observations of Army intelligence analysis (in practice) associated with anticipated requirements for the Unit of Action. We analyze the utility and limitations of a computational model of diagnostic reasoning with respect to intelligence analysis and interpretation and identify classes of knowledge that appear to be essential to performing these tasks. The results are considered with respect to their implications for automated support to intelligence analysis and interpretation.

Introduction

The U.S. Army's Future Combat System, and Future Force, are critically dependent on achieving a level of information superiority that can offset potential vulnerabilities that may arise from lighter (less-hardened) combat vehicles, and to support timely, quality decision-making during high tempo, tactical-level operations often involving an asymmetric threat. The wealth of collection assets envisioned in the Future Force is anticipated to result in unprecedented levels of data requiring analysis and interpretation by intelligence analysts. Data fusion is being considered by some as the solution to handling the severe data overload that analysts can expect to experience. The Joint Directors of Laboratories (JDL) Data Fusion model categorizes data

fusion-related functions at a high level of generality (Steinberg et al, 1998). Solving many different types of problems on all fusion levels is expected to be critical to solving complex analysis and interpretation problems. For fusion in the Army, there has been little published reflecting an elaboration of functionality on levels 2 and 3 of the JDL model, both of which are viewed as critical to solving problems characterizing intelligence analysis and interpretation. Walsh (2002) has offered an initial first-level decomposition of functionality for Level 2. Powell and Broome (2002) and Powell (2002) indicated the complex set of interrelationships among problems within and between fusion levels characterizing Army intelligence analysis and interpretation suggests a user-centric, holistic approach addressing fusion levels 1 through 5 may be most appropriate. The present paper characterizes selected key aspects of analysis and interpretation problems and processes based on observations of Army intelligence analysis (in practice) associated with anticipated requirements for the Unit of Action. We analyze the utility and limitations of a computational model of diagnostic reasoning with respect to intelligence analysis and interpretation and identify classes of knowledge that appear to be essential to performing these tasks. The results are considered with respect to their implications for automated support to intelligence analysis and interpretation at the Unit of Action level.

The work discussed in the present paper is associated with the U.S. Army's Science and Technology Objective Program entitled Fusion-based Knowledge for the Future Force which is addressing operational problems in intelligence analysis and interpretation corresponding to those found on levels 2, 3, 4 and 5 of the JDL Data Fusion Model.

Answering Priority Intelligence Requirements

The volume of reports that must be analyzed today represents a severe information overload on Army analysts at operational-level units. The problem is expected to be exacerbated in Future Force units even at the tactical-level such as the Unit of Action of the Future Combat System.

The present paper focuses on the problem of answering commanders' priority intelligence requirements (PIRs). PIRs are intelligence requirements that are critical to developing, analyzing and executing friendly force (FF) courses of action (COAs). Typically, a given PIR is associated with a single point in a FF COA requiring a decision that needs to be made by the commander. The answer to the PIR will give the commander information critical to making a decision about alternatives that he/she has developed. Other times, a given PIR will not be tied to a particular decision point in a FF COA, but still will be linked to a critical decision; an example might be to find all high-payoff targets in the area of operations so they can either be monitored or targeted.

Answering critical intelligence requirements is an activity present at all echelons. However, the problem we are discussing here occurs at echelons where there is a deliberate and formal process carried out by a staff available to develop PIRs during the military decision making process (MDMP), to analyze those requirements, and to develop ISR tasks to support answering them. In particular, our focus is the Future Combat System's Unit of Action. In these situations, the volume of information that analysts will need to analyze and interpret is expected to significantly strain or exceed their cognitive capacities. Consequently, there is an interest in developing an understanding of this problem with the goal of developing automated support.

When a unit receives an operations order from higher HQ, it includes a set of intelligence preparation of the battlefield (IPB) products, their updated staff estimates of the situation, and initial commander's guidance. Depending on its level in the command hierarchy, the unit will begin carrying out the MDMP either formally (higher levels) or in some abbreviated form (lower levels). Using the order and associated input specified above, the receiving HQ (at the Unit of Action) will carry out a set of tasks called mission analysis. Some of the key tasks the intelligence staff (and often other staff in a supporting role) will perform include developing an initial IPB, determining information requirements and initial PIRs, and determining the initial intelligence, surveillance and reconnaissance plan. The unit's initial IPB product will include enemy situation templates, a modified combined obstacle overlay, and a set of high value targets. Later in the MDMP, during FF COA analysis, the initial set of PIRs may be changed through additions and/or deletions. Each PIR will have a latest time information is of value (LTIOV) associated with it. The LTIOV can be tied to an absolute or relative time, to an event, or to a particular point in the operation. There are no standard sets of PIRs. And the set of PIRs is dynamic, changing with the ebb and flow of the battlespace as the situation evolves pre-engagement, and as the FF COA execution progresses through its phases. The requirements management process is continuous.

Focusing on Tasks (Not on Fusion Levels)

Note that verbal protocols we have collected do not indicate analysts decompose their problems or tasks into different levels of fusion (such as those identified in the JDL Data Fusion model). Instead, the protocols indicate the problem-solving activities characterizing fusion at the different levels appear in virtually all of the functions they must perform. Rather than tackling these intelligence problems by focusing on a particular level of fusion, our work investigates the actual tasks that Unit of Action analysts will need to carry out in fulfilling their responsibilities in analysis, interpretation and intelligence synchronization planning. Our position is that formulating and decomposing the problem in terms of tasks analysts must perform, rather than in terms of fusion levels, will lead to human-machine system designs that are more likely to discover sub-tasks that are good candidates for machine solution, and lead to increased overall human-machine performance.

Using Abstractions of Computational Tasks

A study of doctrine and analysis of verbal protocols collected from Army intelligence analysts (associated with the United States Army Intelligence Center, Fort Huachuca) during the course of answering PIRs in simulated exercises provided a basis for trying to identify the presence of task types that have been framed, in computational terms, as characterizing problem-solving in different domains. In particular, there appear to be several similarities between the problem-solving tasks associated with answering PIRs and those observed in diagnostic reasoning. We have used this set of task types as a model for interpreting the verbal protocols and doctrine with the goal of determining where and to what extent the task types are present when analysts go through the process of answering PIRs.

Diagnostic Reasoning vs. Answering Priority Intelligence Requirements

In the study of problem-solving in some medical domains [e.g., Josephson et al., 1985; Chandrasekaran, 1983; Pople, 1977; Reggia, 1983], protocol data have indicated that clinicians make use of their knowledge of a disease taxonomy to select a disease category based on the data available. Such taxonomies have been developed and evolved over many decades by the medical community. Although we are not ruling this possibility out, our protocol data to date do not suggest such a taxonomy is being used to answer PIRs. If this turns out to be a characteristic of answering PIRs, it may be explained by the context-dependent nature of this problem. Each PIR is tied to a critical decision associated with the FF COA. At a very coarse level of representation, FF COA types may share similarities from one situation to another; however, closer inspection reveals the elements that comprise a given COA will be heavily dependent on the particular FF mission, the enemy, the terrain in the area of operations/interest, the troops (FF units), the time available for planning and acting, and civilians in the area (METT-TC) and how these factors interact with one another. This dependency makes it extremely difficult to generate FF COAs having relevance and utility without knowing the particular METT-TC that they must address; discussions with several analysts indicate METT-TC is unique to each situation. The dependency on FF COA development, analysis and selection on METT-TC also seems to significantly characterize answering PIRs. A number of Army analysts have indicated that the analysis used to decompose a PIR into indicators and specific information requirements (SIRs) is METT-TC dependent. One goal of our ongoing work is to identify aspects of reasoning directed at answering PIRs where there are data-to-evidence and evidence-to-hypothesis mappings that are independent of METT-TC or only weakly dependent (or can be defined broadly enough to be relevant yet narrowly enough such that retrieval and editing could be done in a timely manner). We discuss an example of this later in the paper. In addition, it may be possible to usefully organize knowledge into distinct taxonomies which can be logically connected to each other and used to answer PIRs. More on this topic below.

Diagnostic reasoning has been characterized as a mapping of a set of all subsets of observations of a system to the set of all subsets of possible malfunctions, such that each malfunction subset is a best explanation for the corresponding observation subset (Josephson, 1996). From our analyses, the concept of malfunctions seems to offer little when determining the mapping involved in answering PIRs. Instead, if we think about organizing knowledge into a taxonomic form, it appears that malfunctions should be replaced by a set of other concepts including (minimally) plans, actions, events, units (force structures), and weapons. Each of these concepts would be organized in its own hierarchy. Based on the nature of a given PIR, one or more of these hierarchies would be used to construct an overall answer. Although these hierarchies would need to be related logically, it is not clear at this time what those relations should be exactly. The basic idea is along these lines. Imagine we have a River-Crossing-Operation (RCO) concept in an Actions taxonomy, and we have some evidence to establish it as a concept that we want to refine. Associated with this concept we may have represented knowledge indicating what this action entails such as units (force types) or equipment types. So, we may have pre-defined a relation between this Action taxonomy to the Unit taxonomy that will result in RCO invoking a hypothesis (concept) that Combat-Engineer-Unit should be present (in the area where RCO is hypothesized to be occurring). A plan hierarchy seems to be a good

candidate for integrating concepts comprising the other hierarchies because their domains (actions, units, etc.) can be defined as representing parts of plans.

A number of investigators have aimed to identify task abstractions for different classes of problems [e.g., McDermott, 1988; Chandrasekaran, 1983]. One such abstraction (Josephson, 1996) indicates the elementary task types in diagnosis are:

- hierarchical classification
- hypothesis matching
- knowledge-directed data retrieval
- abductive assembly of hypotheses

Much of the remainder of the present paper uses this particular set of task abstractions, and their associated knowledge requirements, as a framework for attempting to decompose the problem of answering PIRs into computational tasks, and for identifying classes of knowledge that appear to be key to answering commanders' PIRs.

Hierarchical Classification

This has been described as a mapping of a set of observations to the leaf nodes of a classification hierarchy (and including a plausibility or likelihood given the observations).

Figure 1 depicts an example of knowledge encoded in a hierarchy of type-subtype links we have constructed which characterizes tactical operations that may be associated with a particular enemy force. The intent is not to be operationally accurate regarding the particular information in the diagram, but to show how such a taxonomy might be usefully organized as we develop such knowledge about threats. Each node in the hierarchy represents a concept defining offensive operations at different levels of abstraction. Classification would involve identifying a node high in the tree that can be confirmed, ruled out, or suspended based on such knowledge encoded for that concept and the evidence in the current situation that causes the concept to be invoked. Imagine the Attack concept is invoked. If the evidence does not result in the concept being ruled out, or suspended due to inadequate evidence to establish it, control will pass to the successor nodes of Attack where each will be evaluated for rule-out, confirmation or suspension. This cycle is repeated until at least one leaf node of the tree is established such as Fix (subtype of Envelopment Right). If no leaf nodes are established, then no adequate explanation exists given the observed data and the knowledge in the system. If suspension occurs, then classification in a sense is in limbo until additional data result in either a rule-out or an establish. From an Army perspective, the leaf nodes are at the operations task level of the hierarchy.

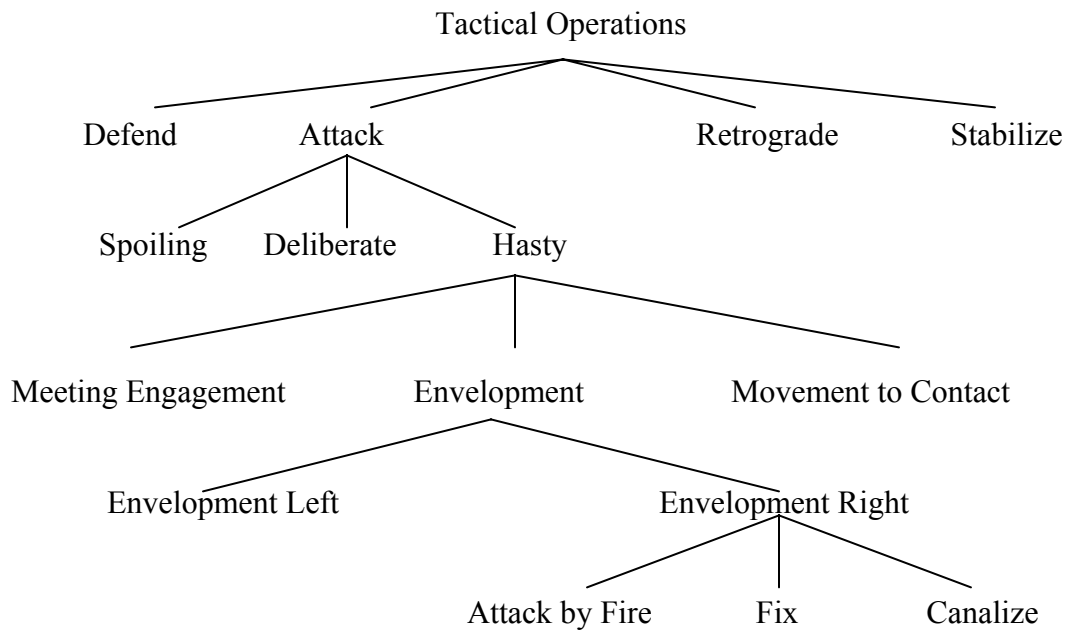


Figure 1. Example Hierarchy of Tactical Operations

Hypothesis Matching

Hypothesis matching involves knowledge representing a hierarchy of abstractions from data to hypotheses. For example, a PIR may be to determine whether a mortar will be present in a particular area (a Named Area of Interest - NAI) during particular phases of FF operations. From verbal protocols, we have identified what appears to be an abstraction hierarchy that could be used, and potentially re-used to a significant extent under any METT-TC, to support answering this PIR. Figure 2 specifies a very small portion of this particular hierarchy, and the concepts represented in it are not modeled at the level of detail they would be for operational use. At the highest level, we may model the concept mortar-emplacement to represent this hypothesis. Subordinate to this level, we may identify a data abstraction level that has concepts: “physical evidence for mortar-emplacement,” “fire-direction-communications evidence for mortar-emplacement,” and “transportation evidence for mortar-emplacement.” Reports of a ring of sandbags, or a round base plate, in the NAI represent data from which “physical evidence for mortar-emplacement” could be inferred. Communications transmissions from that area involving weather-related traffic may provide data supporting an inference that there is “fire-directions communications for mortar emplacement” and so on. An analyst could assign evidence weights to each element of data, and work out a scheme that would combine these weighted pieces of data. Also, the analyst would need to set a threshold for evidence at which a hypothesis would be matched. Although we have protocol data suggesting a particular scheme for assigning weights to data, we do not explore this issue in the present paper. However, we see it as a key issue for developing automated support that will be congruent with approaches analysts use to assign weights.

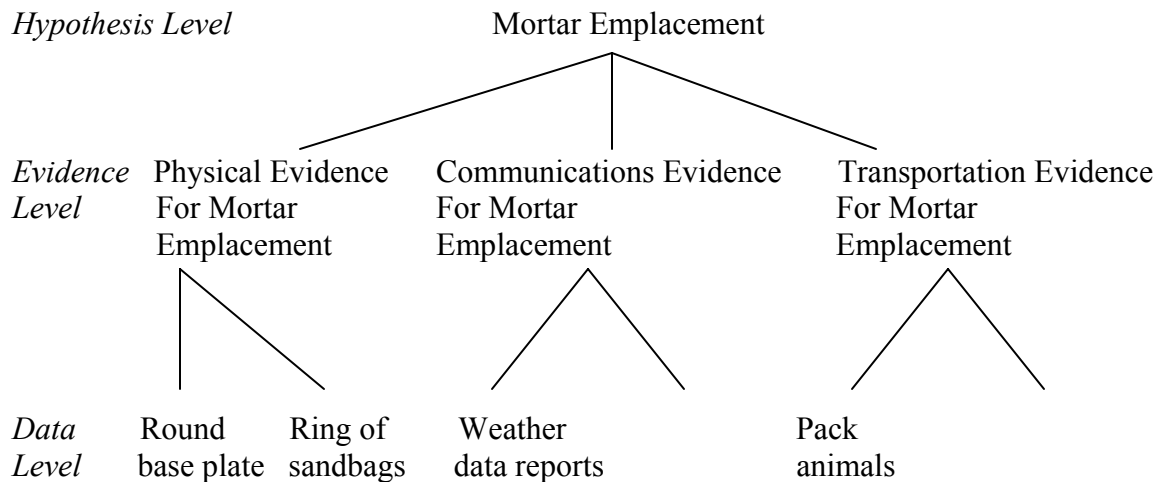


Figure 2. Example of Data-to-Evidence-to-Hypothesis Mappings

Knowledge-directed Data Retrieval

Hypothesis matching needs data in order to operate. In answering PIRs, analysts typically will be flooded with reports of different types that provide data that may be useful. However, there may be times when the analyst will want to check data sources (e.g., use Intelligence Reach) to determine if the required data have already been collected and reported. This may be as straightforward as determining if a particular individual was seen in a particular area during a particular time interval; the answer may be available through querying a data repository. Other times, the situation may be more complex and require knowledge of the relationship between data elements (that may be present in a data repository) and higher-level evidence. Whenever we can identify knowledge of this type that is case-independent, there is an opportunity for re-use across potentially diverse METT-TC situations. For example, an analyst may want to know the cargo capacity of a particular type of truck observed in the area of interest. This kind of knowledge may be represented in a database and derivable, say, via inheritance from the data abstraction representing this class of truck.

Abductive Assembly

Abductive assembly involves assembling a subset of hypotheses into a composite hypothesis that best explains the data. This task involves knowledge of the causal or logical links between hypotheses. Also, knowledge of hypothesis interactions may indicate higher or lower likelihoods for hypotheses appearing together. For example, an analyst may have a hypothesis that the enemy will conduct a river-crossing operation in a particular area during a particular interval because a river cuts through the avenue of approach used by the hypothesized most likely enemy COA. There may also be a hypothesis that an enemy combat engineering unit is advancing toward that area. The mutually supporting relationship between combat engineers and river crossings would increase our estimate of the likelihood of a river crossing in the area. Abductive assembly also requires knowledge of general principles related to concepts such as parsimony that can be used to choose among different hypotheses that explain the same data. Below we

present a further elaboration of classes of knowledge required for Abductive Assembly with respect to using it for answering commanders' PIRs; we illustrate each with an example.

Plausibility-assignment Knowledge. Analysts need to be able to ascribe a plausibility rating to each hypothesis. For example, if sandbags in a ring with a 2-10 meter diameter, and a bipod, are observed in a NAI, analysts should have a means for combining that information such that their presence together would yield a particular likelihood that this is a mortar emplacement.

Explanatory Knowledge. Analysts should be able to specify what a given hypothesis can explain (or account for) about the data. For example, a river crossing hypothesis can explain the presence of bridging equipment on the near side of the river as well as forward observers on the far side of the river.

Hypothesis-hypothesis Interaction Knowledge. This was briefly discussed above in the example of combat engineers and river crossings in a mutually supporting interaction. Explanatory interaction is another type. One form in which this occurs is when a hypothesis is a more detailed refinement of another hypothesis. For example, in Figure 1, an Envelopment is a refinement of a Hasty (Attack). Decisions about how to grow the composite hypothesis would need to be defined in a control mechanism. For example, explanatory coverage (breadth) may be given priority over the level of detail explained in which case Hasty may take precedence over Envelopment. Additional data may reveal Meeting Engagement is more plausible than Envelopment.

Strategy knowledge. This is knowledge required to guide problem-solving; analysts should play a key role in specifying it. The fundamental choice is between a fixed versus an opportunistic control regimen. Verbal protocol data suggest that analysts prefer flexibility in choosing methods to expand an overall explanation. For example, time available to answer a given PIR can influence how much explanatory coverage (breadth) is possible to pursue.

Implications for Automated Assistance and Future Work

Although we have begun to uncover what appear to be certain limitations of diagnostic reasoning as an approach to answering commanders' PIRs, we continue to see potential value in organizing domain knowledge taxonomically for this task, but by factoring it into multiple taxonomies that can be logically linked together. Also, we think the computational abstractions of tasks for diagnosis exhibit merit for decomposing intelligence tasks into relatively well-understood task types. Some of the methods (which we only partially discussed in this paper) for carrying out these tasks show promise for performing tasks involved in answering commanders' PIRs.

The decomposition of the problem into these computational task types also is assisting us in identifying the knowledge types required to perform each task. This addresses the knowledge acquisition bottleneck and is allowing us to develop knowledge acquisition strategies tailored to each task type.

It should be noted that hypothesis generation in this approach is via selection and composition. Selection is partially based on the existence of appropriate taxonomies. One of the criteria for accepting a hypothesis is whether an analyst would consider it to be good enough to explain the data. In cases where our composites are not good enough, we have to hold open the possibility that our knowledge is inadequate (that perhaps the situation is novel, for example).

With an industrial partner, we are about to embark on the development of a proof-of-principle prototype to explore the utility of this computational task abstraction approach for diagnostic reasoning to supporting the development of a software capability to assist in answering commanders' PIRs.

Related Work

Others have attempted to address the problem of answering PIRs using a Bayesian Belief Network (BBN) approach [e.g., Das et al, 2002; Jones et al., 1998]. Although it remains an empirical question, we are optimistic that the approach we are using will be more efficient in terms of knowledge acquisition (discussions with parties participating in these BBN efforts have underscored the significant knowledge acquisition challenges encountered). Our optimism stems from formulating the problem and solution into multiple task types with a tailored knowledge acquisition approach to each task rather than formulating the problem solution within a single representation and inferencing formalism that relies heavily on the human analyst thinking about the problem from a BBN perspective. Also, our approach may offer methods of reasoning that will be more easily understood and accepted by analysts than probabilistic methods. If our solution approach is easier for analysts to understand, it may be easier to validate. These are other issues requiring empirical assessment. We may eventually find particular aspects of this set of task types that will benefit from incorporating a probabilistic approach.

Implications for Training Analysts

Although not a primary goal of our work, we see potential value in the computational models we are developing for task types involved in answering commanders' PIRs, to shed light on useful ways knowledge can be organized, and reasoned with, in carrying out these tasks. If we experimentally show the effectiveness and validity of this approach, it may be worthwhile for the Army Training and Doctrine Command to evaluate it for possible inclusion in the curriculum used to train analysts.

References

- Chandrasekaran, B. Towards a Taxonomy of Problem-solving Types. AI Magazine, Winter/Spring, 9-17, 1983.
- Das, S., Grey, R., and Gonsalves, P. Situation Assessment via Bayesian Belief Networks. Proceedings of the 5th International Conference on Information Fusion, Annapolis, Maryland, July 2002.

Josephson, J.R., Tanner, M.C., Smith, J., Svirebely, J., & Strohm, P. Red: Integrating Generic Tasks to Identify Red-cell Antibodies. In K. N.Karna (Ed.), Proceedings of The Expert Systems in Government Symposium, pp. 524-531, IEEE Computer Society Press, 1985.

Josephson, J.R. and Josephson, S.G.. Abductive Inference: Computation, Philosophy, Technology. Cambridge University Press, 1996.

McDermott, John. Preliminary Steps Toward a Taxonomy of Problem Solving Types. In S. Marcus (Ed.) Automating Knowledge Acquisition for Expert Systems. (pp. 225-256). Boston: Kluwer Academic Publishers, 1988.

Pople, Harry. The Formation of Composite Hypotheses in Diagnostic Problem Solving.: An Exercise in Synthetic Reasoning. In Proceedings of the Fifth International Joint Conference on Artificial Intelligence, pp. 1030-1037, 1977.

Powell, Gerald M. and Broome, Barbara. Fusion-Based Knowledge for the Objective Force. MSS National Symposium on Sensor and Data Fusion, San Diego CA, August 2002.

Powell, Gerald M. Opportunistic Problem-solving in Data Fusion: The Blackboard Model, 23rd Army Science Conference, Orlando FL, December 2002.

Reggia, J.A., Nau, D.S., and Wang, P. Diagnostic Expert Systems Based on a Set Covering Model. International Journal of Man-Machine Studies, 19(5), 437-460, 1983.

Steinberg, Alan., Bowman, Christopher. and White, Franklin. Revisions to the JDL Data Fusion Model, American Institute of Aeronautics and Astronautics, New York, 1998.

Walsh, Dale. Fusion for the Army. Briefing to INSCOM, January 2002